

Some Methodological Uses of Responses to Open Questions and Other Verbatim Comments in Quantitative Surveys

Singer, Eleanor; Couper, Mick P.

Veröffentlichungsversion / Published Version
Zeitschriftenartikel / journal article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:
GESIS - Leibniz-Institut für Sozialwissenschaften

Empfohlene Zitierung / Suggested Citation:

Singer, E., & Couper, M. P. (2017). Some Methodological Uses of Responses to Open Questions and Other Verbatim Comments in Quantitative Surveys. *Methods, data, analyses : a journal for quantitative methods and survey methodology (mda)*, 11(2), 115-134. <https://doi.org/10.12758/mda.2017.01>

Nutzungsbedingungen:

Dieser Text wird unter einer CC BY Lizenz (Namensnennung) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier:
<https://creativecommons.org/licenses/by/4.0/deed.de>

Terms of use:

This document is made available under a CC BY Licence (Attribution). For more Information see:
<https://creativecommons.org/licenses/by/4.0>

Some Methodological Uses of Responses to Open Questions and Other Verbatim Comments in Quantitative Surveys

Eleanor Singer & Mick P. Couper

Survey Research Center, University of Michigan

Abstract

The use of open-ended questions in survey research has a very long history. In this paper, building on the work of Paul F. Lazarsfeld and Howard Schuman, we review the methodological uses of open-ended questions and verbatim responses in surveys. We draw on prior research, our own and that of others, to argue for increasing the use of open-ended questions in quantitative surveys. The addition of open-ended questions – and the capture and analysis of respondents’ verbatim responses to other types of questions – may yield important insights, not only into respondents’ substantive answers, but also into how they understand the questions we ask and arrive at an answer. Adding a limited number of such questions to computerized surveys, whether self- or interviewer-administered, is neither expensive nor time-consuming, and in our experience respondents are quite willing and able to answer such questions.

Keywords: open questions; textual analysis; verbatim comments



© The Author(s) 2017. This is an Open Access article distributed under the terms of the Creative Commons Attribution 3.0 License. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

1 Introduction

More than 75 years ago Lazarsfeld (1935), in “The Art of Asking Why,” offered advice on the proper (and improper) deployment of open-ended questions. He identified six main functions of the open-ended interview: clarifying the meaning of a respondent’s answer, singling out the decisive aspects of an opinion, discovering what has influenced an opinion, determining complex attitude questions, interpreting motivations, and clarifying statistical relationships. In “The Controversy over the Detailed Interview – An Offer for Negotiation,” prepared in response to an invitation to adjudicate professional disagreements over the relative merits of closed versus open-ended questions, he argued that both open and closed questions should be used in a comprehensive research program (Lazarsfeld, 1944).

Over time, the economics of survey research gradually drove out open-ended interviewing as a technique for quantitative large-scale studies (cf. Geer, 1991). But about a quarter century later Howard Schuman proposed an ingenious solution to the cost dilemma. In “The Random Probe” (1966), he pointed out that most of the functions of open-ended questions noted by Lazarsfeld could, in fact, be fulfilled by probing a randomly selected subset of responses to closed-ended questions with open-ended follow-ups. Such probes could be used to clarify reasons for the response, clear up ambiguities, and explore responses that fell outside the expected range of answers. Because they would be put only to a subset of respondents, they would reduce the cost of recording and coding; but since the subsample was randomly selected, the results could be generalized to the sample as a whole. Schuman himself has made much use of this technique over his long career in survey research, reprised in his most recent book, *Meaning and Method* (2008). Nevertheless, the promise of this approach has not yet been fully realized, despite the development of technologies that make it even easier to implement today.

Here, we review several primarily methodological uses of open-ended questions and give examples drawn from our own research as well as that of others. We believe the adaptation of open-ended questions to some functions in quantitative surveys for which they have not previously been used, or used only rarely, will result in more respondent-focused surveys and more accurate and useful data. The paper argues for more inclusion of open-ended questions in quantitative surveys and discusses the technological and methodological advances that facilitate such inclusion. The major advantage of embedding such questions in actual surveys rather than restricting their use to qualitative interviews is the breadth and representativeness of coverage they provide at little additional cost. Such use should

Direct correspondence to

Mick P. Couper, ISR, University of Michigan, P.O. Box 1248, Ann Arbor, MI 48106, U.S.A.

E-mail: mcouper@umich.edu

complement, not replace, the use of open questions and verbatim responses during the instrument development and pretesting process.

We take a broad perspective on open questions in this paper, including any question where the respondent's answers are not limited to a set of predefined response options. Couper, Kennedy, Conrad, and Tourangeau (2011) review different types of such responses, including questions eliciting narrative responses (e.g., "What is the biggest problem facing the country today?") and those soliciting a numeric response (e.g., "During the past 12 months, how many times have you seen or talked with a doctor about your health?"). We include all these types, and expand the notion to include verbatim responses to closed questions that do not fall within the prescribed set of response alternatives.

2 Why Add Open-Ended Questions to Surveys?

As already noted, Schuman (1966) proposed following some closed questions with open-ended probes administered to a random sample of respondents in order to clarify their answers and – which is often forgotten – to establish the validity of closed questions (Schuman & Presser, 1979). We believe such probes can serve a number of other important functions as well. For all of these, embedding the probes in ongoing surveys has clear benefits. First, there is a good chance of capturing the full range of possible responses, since the survey is administered to a random sample of the target population; and second, if the survey is web-based or administered by an interviewer using a computer, the responses can be captured digitally, facilitating automatic transcription or computer-assisted coding, in turn reducing the cost and effort involved in analyzing the responses. Such "random probes" thus provide a useful addition, and in some cases an alternative, to a small number of qualitative interviews administered to convenience samples.

In what follows, we identify seven primarily methodological uses of open-ended questions: Understanding reasons for reluctance or refusal; determining the range of options to be used in closed-ended questions; evaluating how well questions work; testing methodological theories and hypotheses; checking for errors; encouraging more truthful answers; and providing an opportunity for feedback. We omit another frequent use of open-ended questions – namely, as an indicator of response quality (e.g. Galesic & Bosnjak, 2009; for a summary of this use of open-ended questions in incentive experiments see Singer & Kulka, 2002).

2.1 Understanding Reasons for Refusal

The first use of open responses lies outside the traditional domain of standardized survey instruments. Introductory interactions were long thought of as something external to the survey itself, and therefore as something not subject to systematic measurement. However, the early pioneering work of Morton-Williams (1993; see also Morton-Williams & Young, 1987) showed that systematic information can be collected about these interactions and used for quantitative analysis, and a few studies have collected systematic data about “doorstep interactions” between interviewers and respondents in an effort to use respondent comments to predict the likelihood of response and allow interviewers to “tailor” their comments to specific respondent concerns (Morton-Williams & Young, 1987; Morton-Williams, 1993; Groves & Couper, 1996; Campanelli et al., 1997; Couper, 1997; Sturgis & Campanelli, 1998; Groves & McGonagle, 2001; Couper & Groves, 2002; Bates et al., 2008).

In an early paper, Couper (1997) demonstrated that there is some veracity to the reasons sample persons give for not wanting to participate in a survey. Those who say “not interested” did indeed appear to be less interested, engaged, and knowledgeable about the topic (elections) than those (for example) who gave “too busy” as a reason. Interviewer observations are now a standard part of many survey data collection protocols. Often the verbatim reactions of householders to the survey request are field-coded by interviewers. Recent efforts have focused on improving the quality of such observations (see, e.g., West, 2013; West & Kreuter, 2013, 2015).

For example, the US Census Bureau makes data from its contact history instrument (CHI; see, e.g., Tan, 2011), which systematically captures information on interviewer-householder interactions, available to researchers. The CHI provides information about the characteristics of all sample members with whom contact was made, permitting not only the tailoring of subsequent contacts to counteract reservations that may have been expressed at the prior encounter, but also to predict what kinds of responses are likely to lead to final refusals and which are susceptible of conversion. Bates, Dahlhamer, and Singer (2008), for example, analyzed the effect of various respondent concerns, expressed during a personal contact with an interviewer, on cooperation with the National Health Interview Survey. While acknowledging various limitations of the CHI instrument, including the fact that recording and coding the concerns involve subjective judgments by interviewers as well as possible recall error if such concerns are not recorded immediately, the authors report a number of useful findings in need of replication. Thus, for example, although 23.9% of households claimed they were “too busy” to do the interview during at least one contact, 72.8% of households expressing this concern never refused and only 10.3% were final refusals. Similarly, although 13.3% of households expressed privacy concerns, 62.9% of those expressing privacy concerns never

refused, and only 13.9% were final refusals. On the other hand, 34.1% of those (12.7% of households) saying “not interested” and “don’t want to be bothered” never became respondents (*ibid.*, Table 1). Because interactions between interviewers and respondents were not recorded verbatim in this study, we can only surmise why certain concerns were more amenable to mitigation than others, or guess at which interviewer conversational strategies might have been successful. While early methodological studies (most notably Morton-Williams, 1993) had interviewers tape-record the doorstep interactions, most subsequent work has required interviewers to report their observations of the interaction, a process subject to measurement error. Portable, unobtrusive digital recorders, increasingly an integral component of the laptop and tablet computers interviewers are using for data collection, make such doorstep recording increasingly feasible.¹ Recording of introductory interactions in telephone surveys is logistically even easier (e.g., Couper & Groves, 2002; Benki et al., 2011; Conrad et al., 2013).

Modes of interviewing that record the entire interaction, rather than manually recording only the respondent’s concern, could begin to provide answers to questions relating to the process of gaining cooperation. For example, Maynard, Freese, and Schaeffer (2010) draw on conversation-analytic methods and research to analyze interviewer-respondent interactions in order to better understand the process of requesting and obtaining participation in a survey interview. The authors state, “This article contributes to understanding the social action of requesting and specifically how we might use insights from analyses of interaction to increase cooperation with requests to participate in surveys.” Or, as the authors of the CHI paper note, “The potential of these new data to expand our understanding of survey participation seems great since they are collected at every contact, across modes, and across several different demographic surveys for which the US Census Bureau is the collecting agent.” Indeed, they include an analysis of Consumer Expenditure Survey Data that replicates key findings of the main analysis (Bates et al., 2008).

2.2 Determining the Range of Options to Be Offered in Closed-Ended Questions

In “The Open and Closed Question”, Schuman and Presser (1979) talk about the two main functions of open-ended questions: Making sure that all possible response options are included in the final questionnaire, and avoiding bias. They investigate experimentally how closely the coding of responses to an open-ended question replicates the *a priori* response alternatives assigned to a question about the importance of different aspects of work. Schuman has also talked about the

1 Note, however, that the technical developments do not address the informed consent issues raised by recording such introductory interactions.

importance of ascertaining the full range of response options to controversial questions before constructing a questionnaire. What, for example, is the most extreme response option to a question about the conditions under which abortion should be forbidden? Is it the termination of any pregnancy, however brief, or does it extend to the prevention of conception after unprotected intercourse, or even to the use of contraception? Schuman has suggested talking to groups holding extreme positions on both sides of a controversial issue before drafting questions about it. A possibly attractive alternative is to include the question in open-ended form – e.g., “What kinds of actions would you include in a definition of abortion?” – on a survey of a random sample of the target population which precedes the planned survey on abortion attitudes. Such a question should yield not only the extremes but also a distribution of intermediate responses. This is analogous to doing a small number of qualitative, semi-structured interviews prior to fielding a questionnaire, but has the advantage of doing so with a larger, more diverse sample in an ongoing survey at marginal cost. Behr et al. (2012, 2013, 2014) have investigated some factors contributing to the success of such probes in web surveys.

2.3 Evaluating How Well Questions Work

Just as open questions administered to a random sample can be useful in developing a questionnaire, so they can be useful in evaluating how well questions work in an actual survey. Martin (2004) discusses at length the use of open and closed debriefing questions administered after the main survey for evaluating respondents’ understanding of key questions. Such questions have been used to measure the accuracy of respondents’ interpretation of terminology, questions, or instructions; to gauge respondents’ reactions or thoughts during questioning; and to obtain direct measures of missed or misreported information (e.g. Belson, 1981; DeMaio, 1983; DeMaio & Rothgeb, 1996; Oksenberg et al., 1991; Schuman, 1966). Hess and Singer (1995), for example, used open as well as closed questions administered to a random subsample of respondents to see how well respondents understood questions on a Food Insecurity supplement and how reliably some questions were answered.

Given the increasing ease with which digital recordings of the entire interview can be captured for analysis, verbatim responses to closed-ended questions in interviewer-administered surveys are becoming increasingly useful for evaluating the performance of survey questions. In the days of paper-and-pencil surveys, interviewers recorded the interviews on tape recorders. These were painstakingly coded and analyzed using methods such as behavior coding (see, e.g., Fowler & Cannell, 1996) or conversational-analytic methods (e.g., Schaeffer & Maynard, 1996; Maynard et al., 2002), often only in small pretests. Digital recordings integrated into computer-assisted interviewing (CAI) software make the task of finding responses to specific questions much easier. While much of the focus of this work has been on

evaluating interviewers, we believe such recordings are a valuable tool for evaluating survey questions. Indeed, Cannell and Oksenberg (1988) identified three main objectives of interview observation: 1) to monitor interviewer performance, 2) to identify survey questions that cause problems for the interviewer or respondent, and 3) to provide basic data for methodological studies.

To give one recent example: in the process of developing an online version of the Health and Retirement Study (see <http://hrsonline.isr.umich.edu/>) instrument, we were struggling with how to refer to family members (siblings or children) who had died since the last wave of data collection. HRS staff selected a number of recordings from the prior interviewer-administered wave of the survey where the data revealed a death of a sibling or child. By listening to these interactions, they were able to determine that the term “deceased” was used more frequently than “passed (away)” or other terms when referring to such family members. This enabled us to recommend appropriate wording for the online version of the survey.

Other examples of such targeted analysis include identifying questions with high rates of missing data to understand how respondents are communicating their responses; identifying concerns expressed about in-survey consent requests; understanding how respondents might qualify their answers in response to questions asking for exact qualities (e.g., income or assets, life expectancy probability, etc.); and the like. Both survey data and paradata can be used to identify questions for more detailed examination, whether qualitative or quantitative. We believe this is an under-utilized opportunity to use existing digital recordings to evaluate and improve survey questions.

2.4 Testing Methodological Theories and Hypotheses

Porst and von Briel (1995), Singer (2003), and Couper et al. (2008, 2010) have used open-ended questions in face-to-face, telephone, and online surveys to explore reasons people give for being willing (or unwilling) to participate in a hypothetical survey. Those who said they would be willing to participate cited things like wanting their opinions to be heard or wanting to contribute to the research goals, or their interest in the topic of the survey or the incentive associated with participation. Those who said they would not be willing to participate gave some general reasons – not interested, too long, too little time – as well as a large number of responses that were classified as privacy-related (e.g., Don’t like intrusions; don’t like to give financial information). A large number of responses pertained to survey characteristics, such as the topic or the sponsor, and a small number of comments indicated that respondents did not view the survey as offering enough benefits to make participation worthwhile.

These reasons can be reliably coded into a relatively small number of general categories – an egoistic-altruistic dimension (for example, “For the money,” “To

help with the research”), another having to do with situational characteristics (for example, “I’m too busy,” “I’m retired, so I have the time”), and still others having to do with characteristics of the survey (“It’s too long,” “I trust the sponsor”). Such categories could be used to develop a set of exhaustive, mutually exclusive reasons for (non)response, which in turn could be used to test hypotheses or theories about survey participation (Singer, 2011).

We have also asked respondents whether they would, or would not, be willing to permit researchers to make use of paradata – data automatically produced as a byproduct of answering survey questions on web-based surveys – both in connection with hypothetical vignettes and after completing an actual online survey (Couper & Singer, 2013; Singer & Couper, 2011), and followed this with open-ended questions about the reasons for their response. Exploratory questions about whether, and why, respondents would forbid or allow the use of paradata helped clarify the experimental results and can serve as the basis for subsequent quantitative surveys. For example, although we explained to respondents that we never track their browsing behavior, a large number of answers to open-ended questions referred to concerns about tampering with the respondent’s computer, making clear that we had failed to reassure respondents on this point. Subsequent studies could test whether alternative reassuring messages are capable of reducing these concerns and increasing rates of participation. Recording and analyzing the responses given when respondents are asked for consent to linkage to administrative records (e.g., Sakshaug et al., 2012) or for physical or biomedical measurement (e.g., Sakshaug et al., 2010) could similarly help to identify and address reasons for non-compliance.

Examples also exist in other domains of the use of open-ended questions to aid in testing substantive or methodological hypotheses (our focus here being on the latter). For example, Yan, Curtin, and Jans (2010) used an open-ended question on income to measure trends in item nonresponse, which they hypothesized as being inversely related to trends in unit nonresponse. Mason, Carlson, and Tourangeau (1994) used an open-ended question to clarify the subtraction effect in answering part-whole questions. Tourangeau and colleagues (2014, 2016) used open-ended questions to understand the effect of using examples in survey questions.

2.5 Some Other Uses for Open-Ended Questions

In addition to those just discussed, we have found three other uses for open-ended questions. One relatively trivial use is as a check on the coding of the closed question that precedes the open-ended probe. In one particularly dramatic example drawn from our own research (Couper et al., 2008, 2010) we discovered, as a result of working with the open-ended responses, that the codes for answers to the question about willingness to participate had been reversed: Those who had said they would be willing to participate had been coded as if they would refuse, and vice

versa. Less dramatic examples occur more frequently: Someone who checks “9”, indicating great willingness to participate, then enters a response to the open-ended probe that indicates the reverse – for example, “I probably wouldn’t answer these kinds of questions in a face-to-face interview.” It is then possible to correct the response to the closed question or, if the correct coding is not obvious, omit it altogether. Though most of the time they may not be worth the extra effort required, such checks can help to uncover problems with the closed question preceding the probe, and if even small errors cannot be tolerated, the effort may well prove worthwhile.

Still another function of open-ended questions appears to be to permit respondents to give more socially undesirable answers to threatening questions. This function was already pointed out by Blair et al. (1977) with respect to reports about sensitive behaviors such as the amount of alcohol drunk and the frequency of sexual intercourse. Compared with closed questions, open-ended questions elicited reports of a greater average number of drinks and more frequent sexual behaviors, whereas reports about non-sensitive behaviors, such as participation in sports, were not affected by the form of the question. In a subsequent study Tourangeau and Smith (1996) found that “responses to open-ended questions generally fell between responses to the two closed versions,” one of which had response options emphasizing the low end of the sex partner distribution, the other emphasizing the high end. Dykema and Schaeffer (n.d.), reanalyzing the original study by Blair et al. plus additional experiments, concluded that “while closed questions result in higher reports of occurrence, the means among those engaging in the behaviors are usually greater with open questions.” They attribute the difference in means to three factors: the composition of the sample, which is affected by whether a filter question is used; more frequent reporting of high frequencies with open questions; and whether those who report never engaging in the behavior are included in the analysis. In fact, they find that “open questions produce higher estimates of means for nonthreatening as well as threatening behaviors, and do not always do so for threatening questions (p. 24).”

Our research appears to have uncovered another version of this effect on the reporting of socially undesirable feelings. In research exploring race of interviewer effects using real and virtual interviewers, Krysan and Couper (2003) found some cases where white respondents (for example) gave more negative responses to live interviewers than to virtual ones. In qualitative debriefings of respondents, some mentioned that talking to an interviewer gave them an opportunity to explain their choice of responses; in the virtual interviewer condition (as on the web; see Krysan & Couper, 2005), they could only pick one of the response options provided, without the opportunity to justify their choice. Building on this observation, Couper (2012) conducted a web-based experiment in the Netherlands, using a series of questions on attitudes towards immigrants. Half the respondents were given the closed-ended

items, while the other half were presented both the closed-ended responses and an open-ended text box in which they could (if they wished) offer an explanation for their choice of responses. Surprisingly, offering such an option was associated with significantly more *positive* views towards immigrants. One alternative explanation is that the added open question encourages deeper cognitive processing of the question (i.e., thinking of reasons for or against endorsing the statement), potentially leading to more moderate views. Clearly, this finding suggests more research is needed on the role that optional open questions may play in the response process.

It is common in web surveys to limit respondents to one of the available options. In paper surveys, no such restrictions can be made (see Couper, 2008), and it is not uncommon for respondents to avail themselves of the opportunity to add additional information. Similarly, in interviewer-administered surveys, respondents often qualify their answers, express uncertainty, and the like. Much of this information is ignored in the interviewer's entry of the responses into the computer or in the keying of paper questionnaires. Automatic recording of the verbatim responses makes Schuman's (1966) idea of the random probe much more feasible, both for substantive and for methodological purposes. Adding such probes in web surveys, as Behr and her colleagues (2012, 2013, 2014) have shown, is relatively easy. If responses to such follow-up questions are not required, this is unlikely to have a negative effect on survey response.

A final use of open-ended questions is the "anything else" question sometimes appended to a structured questionnaire or interview: "Is there anything else you would like to tell us?" or "Are there any other comments you would like to make?" This is a write-in with a large text box in self-completion surveys (whether paper or web), or an open-ended question where the interviewer is supposed to record the answer verbatim, in interviewer-administered surveys. Often such responses are ignored or – at best – briefly scanned for key concerns, but rarely systematically coded and analyzed. Such a question may help give voice to respondents and may in turn provide us with valuable information, provided we make use of the information contained in the responses. As Peter Lynn has suggested (personal communication), such questions could be used to determine whether their inclusion affects response rates or panel attrition, or whether they affect related matters, such as respondent satisfaction with the survey.

3 Technological Developments Facilitating the Use of Open-Ended Questions

We've already noted that recent technical developments are facilitating both the capture and analysis of open-ended responses. In paper-based surveys, interviewers were either expected to transcribe the respondent's verbatim answers to open

questions, or to record them for later transcription, coding, and/or analysis. With the advent of computer-assisted interviewing, early concerns about requiring interviewers to type such responses into the computer proved to be largely unfounded (see, e.g., Bernard, 1989; Catlin & Ingram, 1988), but the introduction of CAPI may have served to accelerate the decline in the use of open-ended questions.

Field interviewers were often required to carry tape recorders to record entire interviews for quality control or methodological research purposes. Despite the intrusiveness of these devices, respondent consent rates to recording interviews were relatively high (see, e.g., Dykema et al., 1997; see also McGonagle et al., 2015). However, the equipment presented logistical difficulties for interviewers, both during and after the interview. Administrative effort was associated with labeling the cassettes and mailing them to a central office. Confidentiality concerns were raised regarding the handling and storage of the physical media. Analog cassette tapes also presented a big hurdle for coding and analysis. Coders had to search through the tapes to find the relevant sections, or be forced to listen to the entire interview. While the logistical and administrative procedures were somewhat less onerous in telephone surveys (e.g., larger keyboards made typing of open-ended responses easier, centralization facilitated the handling of recording equipment and cassettes), analysis of the responses or coding of the recordings remained a burdensome activity, and one that was hard to do selectively.

The development of digital recording devices made the capture of open-ended responses – and indeed the verbatim responses to all questions in the survey – much easier. Almost all laptop or tablet computers have the built-in capability for digital recording, obviating the need to carry additional equipment. Further, such recording can easily be integrated into the computer assisted interviewing (CAI) software (e.g., Thissen et al., 2013). Indeed, this tool, now known as computer-assisted recorded interviewing (CARI; see Arceneaux, 2007; Hicks et al., 2010; Thissen, 2014; Thissen et al., 2013) is a standard feature of some CAI systems. This brings several benefits for capture: 1) no need for additional equipment (although some laptop microphones are not ideal for recording both interviewers and respondents; see Hansen et al., 2005); 2) the consent process can be automated as part of the survey instrument (recording is automatically activated upon consent); 3) selected parts of the survey (sections or individual items) can be recorded; 4) sound files can be encrypted and transmitted to the central office as part of the regular send/receive activities; and 5) sound files can be easily identified (e.g., by sample ID and question number), facilitating the task of finding particular questions to listen to, transcribe and/or code. Although much of the work using CARI has focused on evaluating interviewer performance, we believe the tool also has great promise for revealing what respondents are saying – and how they are saying it, in response to both to open-ended and (ostensibly) closed-ended questions.

In similar fashion, the increased use of web surveys (relative to paper) makes the task of capturing typed responses to open-ended questions easier and more amenable to analysis. While adding such questions to web surveys may increase the perceived burden on respondents, making such questions optional may reduce this effect. Further, using randomization (as envisioned by Schuman, 1966) could further mitigate any negative consequences. Giving respondents an option to voice their own opinions may even have positive consequences, although this is largely untested (as we discuss elsewhere). Analysis of these responses from web surveys is facilitated by the fact that they are already in digital form.

Turning to analysis, a number of recent developments have made the analysis of open-ended data a much more tractable task. Specifically, recent improvements to several software packages for qualitative analysis make them more useful for the analysis of responses to open-ended questions (see Hughes, 2011). Further information on developments in the area of computer assisted qualitative data analysis (CAQDAS) can be found at the website <http://www.surrey.ac.uk/sociology/research/researchcentres/caqdas/support/analysingsurvey/>.

While not replacing the role of the researcher in developing and identifying themes for coding, these packages facilitate the task of coding itself.

The rapid development of software tools to facilitate the coding and analysis of textual materials in social media – whether through text mining or text analytics or more straightforward sentiment analysis – is expanding the opportunity for researchers to make use of fully-automated or semi-automated processes for coding of open text (see, e.g., Shonlau & Couper, 2016; Klochikhin & Boyd-Graber, 2017). Generic text analytic software, such as the natural language toolkit for Python (www.nltk.org; see also Bird et al., 2009) further facilitates the task of analysis. While these tools have not yet been widely embraced by survey researchers, and further exploration and evaluation is needed, they offer great promise for making the analysis of open-ended question less costly and time-consuming.

Another area of promising development lies in software to convert recorded speech to text. While such speech recognition software might not be ready for the task of converting large numbers of short segments to text (most systems require extensive training to improve recognition for a single user), they can potentially assist in substantially reducing the burden of manual transcription that is necessary for computer-assisted analysis of qualitative data. Recent advances in speech recognition, along with the development of powerful software tools to facilitate coding of text, promise to change the cost and effort equation for dealing with responses to open-ended questions.

4 Discussion

Instead of simply forcing respondents to agree (or otherwise) with the statements we proffer, or pick one of the responses we provide, we can give them an opportunity to tell us what's on their mind with respect to the topic under discussion – whether by offering an explicit open-ended question or by capturing everything they say during the interview. Wenemark (2010) suggests that this may empower and motivate respondents, and O'Cathain and Thomas (2004) go further in suggesting that open questions may help redress the power imbalance between researchers and respondents. However, this in turn obliges us to *listen* to what they say or *read* what they write.

We live in the digital age, where textual responses are readily analyzable using powerful text-analytic software, and where digital recordings of oral responses are increasingly amenable to automatic transcription. The cost of capturing this additional information has been dramatically reduced, and the ease with which it can be coded and analyzed has greatly increased. Yet we still seem to be operating as if paying attention to what respondents say – and the way they say it – is too costly and time-consuming for quantitative study.

The primary barriers to including open-ended responses in questionnaires or capturing verbatim responses relate to 1) concerns about lengthening the interview, 2) the risk of digression, 3) relying on interviewers faithfully recording the information, and 4) the cost of transcribing, coding and analyzing the resulting data. We address each of these objections briefly in turn.

The first two concerns are related. By encouraging respondents to provide open-ended responses, it is believed that interview length is increased and that “bad” respondent behavior is encouraged. Similarly, if interviewers are seen writing down everything that respondents say, this may encourage digression. While legitimate, these concerns are often taken to the extreme, leading to an avoidance of *any* open questions. By capturing responses unobtrusively, we reduce the risk of digression, and need to rely less on interviewers to record the responses as accurately as possible. Having interviewers paraphrase the respondents' answers to open-ended questions may still be valuable, but this could easily be supplemented with the actual words used by the respondents. Giving respondents an opportunity to voice their own views in their own words on key topics covered in the survey may well increase respondent engagement in the interview. This may be especially valuable in panel surveys, where cooperation in later waves is an important consideration.

The costs of processing and analyzing the open-ended responses remain a key concern. Recent software developments have made this a less-onerous undertaking, but it still requires effort. However, with digital recording, analysis can be done selectively, focusing on key questions identified prior to the start of data collec-

tion (e.g., those subject to random probes) or identified during data collection (e.g., by using paradata analysis of keystrokes to identify potentially problematic items). Further, selected subsets of interviews can be analyzed, potentially focusing on key subgroups of interest, such as those who provided a particular type of response or those who gave an indication of having difficulty with the question (again, as revealed through paradata; see Couper et al., 1997; Couper & Kreuter, 2011). In other words, technology has made it much easier to identify selected segments of an interview, and to identify subsets of interviews, questions, or respondents for more intensive analysis, reducing the effort and expense of such work.

As we have said earlier, we are not advocating a return to the days of unstructured interviews. Rather, we are arguing for the judicious use of open-ended questions to support the methodological goals outlined earlier. The verbatim responses we get to closed-ended questions, long ignored by survey researchers, may open up whole new areas of important methodological inquiry, providing valuable insights into the meaning and quality of the information respondents are providing as well as their motivation (or lack of it) for doing so.

We believe that the time has come to give greater voice to respondents in standardized surveys – to give them an opportunity, within the constraints of a structured interview, to express their views on the topics addressed in the survey in their own words. This is relevant to both interviewer-administered and self-administered surveys. Opening up the standardized survey in this way can be of benefit both to respondents (giving them a greater sense of engagement in the interaction) and to researchers (giving us more richly textured data on the topics we are studying and providing methodological insights into the process itself). Technological developments have facilitated this change, but inertia has inhibited us from using them to achieve these goals.

References

- Arceneaux, T. A. (2007). Evaluating the Computer Audio-Recorded Interviewing (CARI) Household Wellness Study (HWS) Field Test. *Proceedings of the Survey Research Methods Section, American Statistical Association*, 2811-2818.
- Bates, N. A., Dahlhamer, J., & Singer, E. (2008). Privacy Concerns, Too Busy, or Just Not Interested: Using Doorstep Concerns to Predict Survey Nonresponse. *Journal of Official Statistics*, 24(4), 591-612.
- Behr, D., Kaczmirek, L., Bandilla, W., & Braun, M. (2011). Asking Probing Questions in Web Surveys: Which Factors Have an Impact on the Quality of Responses? Mannheim: GESIS working paper.
- Behr, D., Braun, M., Kaczmirek, L., & Bandilla, W. (2012). Asking Probing Questions in Web Surveys: Which Factors have an Impact on the Quality of Responses. *Social Science Computer Review*, 30(4), 487-498.

- Behr, D., Kaczmirek, L., Bandilla, W., & Braun, M. (2013). Testing the Validity of Gender Ideology Items by Implementing Probing Questions in Web Surveys. *Field Methods*, 25(2), 124-141.
- Behr, D., Bandilla, W., Kaczmirek, L., & Braun, M. (2014). Cognitive Probes in Web Surveys: On the Effect of Different Text Box Size and Probing Exposure on Response Quality. *Social Science Computer Review*, 32(4), 524-533
- Benkí, J. R., Broome, J., Conrad, F. G., Groves, G. R., & Kreuter, F. (2011). Effects of Speech Rate, Pitch, and Pausing on Survey Participation Decisions. Paper presented at the annual conference of the American Association for Public Opinion Research, Phoenix, May.
- Belson, W. A. (1981). *The Design and Understanding of Survey Questions*. Aldershot, UK: Gower.
- Bernard, C. (1989). Survey Data Collection Using Laptop Computers. Paris: INSEE (Report No. 01/C520).
- Bird, S., Klein, E., & Loper, E. (2009). *Natural Language Processing with Python: Analyzing Text with the Natural Language Toolkit*. Boston: O'Reilly Media.
- Blair, E., Sudman, S., Bradburn, N. M., & Stocking, C. (1977). How to Ask Questions about Drinking and Sex: Response Effects in Measuring Consumer Behavior. *Journal of Marketing Research*, 14, 316-321.
- Campanelli, P., Sturgis, P., & Purdon, S. (1997). Can You Hear Me Knocking: An Investigation into the Impact of Interviewers on Survey Response. London: Social and Community Planning Research.
- Cannell, C. F. & Oksenberg, L. (1988). Observation of Behavior in Telephone Interviews. In R. M. Groves, P. P. Biemer, L. E. Lyberg, J. T. Massey, W. L. Nicholls II, & J. Waksberg (eds.), *Telephone Survey Methodology* (pp. 475-495). New York: Wiley.
- Catlin, G., & Ingram, S. (1988). The Effects of CATI on Costs and Data Quality: A Comparison of CATI and Paper Methods in Centralized Interviewing. In R. M. Groves, P. P. Biemer, L. E. Lyberg, J. T. Massey, W. L. Nicholls II, & J. Waksberg (eds.), *Telephone Survey Methodology* (pp. 437-450). New York: Wiley.
- Conrad, F. G., Broome, J. S., Benki, J. R., Kreuter, F., Groves, R. M., Vannette, D., McClain, C., et al. (2013). Interviewer Speech and the Success of Survey Invitations. *Journal of the Royal Statistical Society, Series A*, 176(1), 191-210.
- Couper, M. P. (1997). Survey Introductions and Data Quality. *Public Opinion Quarterly*, 61(2), 317-338.
- Couper, M. P. (2008). Technology and the Survey Interview/Questionnaire. In M. F. Schober & F. G. Conrad (eds.), *Envisioning the Survey Interview of the Future* (pp. 58-76). New York: Wiley.
- Couper, M. P. (2012). Reducing the Threat of Sensitive Questions in Online Surveys. Paper under preparation for presentation at the General Online Research Conference, Mannheim, March.
- Couper, M. P., & Groves, R. M. (2002). Introductory Interactions in Telephone Surveys and Nonresponse. In D. W. Maynard, H. Houtkoop-Steenstra, N. C. Schaeffer, & J. van der Zouwen (eds.), *Standardization and Tacit Knowledge: Interaction and Practice in the Survey Interview* (pp. 161-177). New York: Wiley.
- Couper, M. P., Hansen, S. E., & Sadosky, S. A.. (1997). Evaluating Interviewer Performance in a CAPI Survey. In L. E. Lyberg et al. (eds.), *Survey Measurement and Process Quality* (pp. 267-285). New York: Wiley.

- Couper, M. P., Kennedy, C., Conrad, F. G., & Tourangeau, R. (2010). Designing Input Fields for Non-Narrative Open-Ended Responses in Web Surveys. *Journal of Official Statistics*, 27(1), 65-85.
- Couper, M. P., & Singer, E. (2013). Informed Consent for Web Paradata Use. *Survey Research Methods*, 7(1), 57-67.
- Couper, M. P., Singer, E., Conrad, F. G., & Groves, R. M. (2008). Risk of Disclosure, Perceptions of Risk, and Concerns about Privacy and Confidentiality as Factors in Survey Participation. *Journal of Official Statistics*, 24, 255-275.
- Couper, M. P., Singer, E., Conrad, F. G., & Groves, R. M. (2010). Experimental Studies of Disclosure Risk, Disclosure Harm, Topic Sensitivity, and Survey Participation. *Journal of Official Statistics*, 26(2), 287-300.
- DeMaio, T. J. (1983). *Approaches to Developing Questionnaires*, Statistical Policy Working Paper 10, Subcommittee on Questionnaire Design, Federal Committee on Statistical Methodology. Washington, DC: Office of Management and Budget.
- DeMaio, T. J., & Rothgeb, J. (1996). Cognitive Interviewing Techniques: In the Lab and in the Field. In Norbert Schwarz and Seymour Sudman (eds.), *Answering Questions: Methodology for Determining Cognitive and Communicative Processes in Survey Research* (pp. 177-195). San Francisco: Jossey-Bass.
- Dykema, J., Lepkowski, J. M., & Blixt, S. (1997). The Effect of Interviewer and Respondent Behavior on Data Quality: Analysis of Interaction Coding in a Validation Study. In P. P. Biemer, R. M. Groves, L. E. Lyberg, N. A. Mathiowetz, & S. Sudman (eds.), *Measurement Errors in Surveys* (pp. 287-310). New York: Wiley.
- Dykema, J., & Schaeffer, N. C. (n.d.). Effects of Open, Closed, and Filter Questions on Measuring the Occurrence and Frequency of Behaviors. Unpublished ms., University of Wisconsin, Madison.
- Fowler, F. J., & Cannell, C. F. (1996). Using Behavioral Coding to Identify Cognitive Problems with Survey Questions. In N. Schwarz & S. Sudman (eds.), *Answering Questions: Methodology for Determining Cognitive and Communicative Processes in Survey Research* (pp. 15-36). San Francisco: Jossey-Bass.
- Fuchs, M. (2009). Asking for Numbers and Quantities: Visual Design Effects in Paper and Pencil Surveys. *International Journal of Opinion Research*, 21(1), 65-84.
- Galesic, M., & Bosnjak, M. (2009). The Effects of Questionnaire Length on Participation and Indicators of Response Quality in a Web Survey. *Public Opinion Quarterly*, 73(2), 349-360.
- Geer, J. G. (1991). Do Open-Ended Questions Measure Salient Issues? *Public Opinion Quarterly*, 55(3), 360-370.
- Groves, R. M., & Couper, M. P. (1996). Contact-Level Influences on Cooperation in Face-to-Face Surveys. *Journal of Official Statistics*, 12(1), 63-83.
- Groves, R. M., & McGonagle, K. A. (2001). A Theory-Guided Interviewer Training Protocol Regarding Survey Participation. *Journal of Official Statistics*, 17(2), 249-266.
- Hansen, S. E., Krysan, M., & Couper, M. P. (2005). Sound Bytes: Capturing Audio in Survey Interviews. Poster presented at the annual meeting of the American Association for Public Opinion Research, Miami Beach, May.
- Hicks, W. D., Edwards, B., Tourangeau, K., McBride, B., Harris-Kojetin, L. D., & Moss, A. J. (2010). Using CARI Tools to Understand Measurement Error. *Public Opinion Quarterly*, 74(5), 985-1003.

- Hughes, G. (2011). Using Qualitative Software to Analyse Open-Ended Survey Questions: A Guide for Quantitative Analysts. Paper presented at the European Survey Research Association Conference, Lausanne, July.
- Hess, J. C., & Singer, E. (1995). The Role of Respondent Debriefing Questions in Questionnaire Development. *Statistical Research Division Working Papers in Survey Methodology (#95-18)*. U.S. Census Bureau. Available online at <http://www.census.gov/srd/papers/pdf/sm9518.pdf>.
- Israel, G. D. (2010). The Effects of Answer Space Size on Responses to Open-Ended Questions in Mail Surveys. *Journal of Official Statistics*, 26(2), 271-285.
- Klochikhin, E., & Boyd-Graber, J. (2017). Text Analysis. In I. Foster, R. Ghani, R. S. Jarmmin, F. Kreuter, & J. Lane (eds.), *Big Data and Social Science: A Practical Guide to Methods and Tools* (pp. 187-214). Boca Taron, FL: CRC Press.
- Krysan, M., & Couper, M. P. (2003). Race in the Live and Virtual Interview: Racial Defiance, Social Desirability, and Activation Effects in Attitude Surveys. *Social Psychology Quarterly*, 66(4), 364-383.
- Krysan, M., & Couper, M. P. (2005). Race-of-Interviewer Effects: What Happens on the Web? *International Journal of Internet Science*, 1(1), 5-16.
- Lazarsfeld, P. F. (1935). The Art of Asking Why. *National Marketing Review*, 1, 32-43.
- Lazarsfeld, P. F. (1944). The Controversy over Detailed Interviews: An Offer for Negotiation. *Public Opinion Quarterly*, 8, 38-60.
- Martin, E. (2004). Vignettes and Respondent Debriefing for Questionnaire Design and Evaluation. In S. Presser, J. Rothgeb, M. P. Couper, J. T. Lessler, E. A. Martin, J. Martin, & E. Singer (eds.), *Methods for Testing and Evaluating Survey Questionnaires* (pp. 149-171). New York: Wiley.
- Mason, R., Carlson, J. E., & Tourangeau, R. (1994). Contrast Effects and Subtraction in Part-Whole Questions. *Public Opinion Quarterly*, 58(4), 569-578.
- Maynard, D., Freese, J., & Schaeffer, N. C. (2010). Calling for Participation: Requests, Blocking Moves, and Rational (Inter)Action in Survey Introductions. *American Sociological Review*, 75(5), 795-814.
- Maynard, D., Houtkoop-Steenstra, H., Schaeffer, N. C., & van der Zouwen, J. (eds.). (2002). *Standardization and Tacit Knowledge; Interaction and Practice in the Survey Interview*. New York: Wiley.
- McGonagle, K. A., Brown, C., & Schoeni, R. F. (2015). The Effects of Respondents' Consent to Be Recorded on Interview Length and Data Quality in a National Panel Study. *Field Methods*, 27(4), 373-390.
- Morton-Williams, J. (1993). *Interviewer Approaches*. Aldershot: Dartmouth.
- Morton-Williams, J., & Young, P. (1987). Obtaining the Survey Interview – An Analysis of Tape-Recorded Doorstep Introductions. *Journal of the Market Research Society*, 29(1), 35-54.
- O'Cathain, A., & Thomas, K. J. (2004). 'Any Other Comments?' Open Questions on Questionnaires – A Bane or a Bonus to Research? *BMC Medical Research Methodology*, 4(25).
- Oksenberg, L., Cannell, C. F., & Kalton, G. (1991). New Strategies for Pretesting Survey Questions. *Journal of Official Statistics*, 7, 349-356.
- Porst, R. & von Briel, C. (1995). Wären Sie vielleicht bereit, sich gegebenenfalls noch einmal befragen zu lassen? Oder: Gründe für die Teilnahme an Panelbefragungen. ZUMA-Arbeitsbericht, Nr. 95/04, Mannheim, Germany.

- Presser, S., Rothgeb, J., Couper, M. P., Lessler, J. T., Martin, E., Martin, J., & Singer, E. (eds.). (2004). *Methods for Testing and Evaluating Survey Questionnaires*. New York: Wiley.
- Sakshaug, J. W., Couper, M. P., & Ofstedal, M. B.. (2010). Characteristics of Physical Measurement Consent in a Population-Based Survey of Older Adults. *Medical Care*, 48(1), 64-71.
- Sakshaug, J. W., Couper, M. P., Ofstedal, M. B., & Weir, D. (2012). Linking Survey and Administrative Records: Mechanisms of Consent. *Sociological Methods and Research*, 41(4), 535-569.
- Schaeffer, N. C., & Maynard, D. W.. (1996). From Paradigm to Prototype and Back Again: interactive Aspects of Cognitive Processing in Standardized Survey Interviews. In N. Schwarz & S. Sudman (eds.), *Answering Questions: Methodology for Determining Cognitive and Communicative Processes in Survey Research* (pp. 65-88). San Francisco: Jossey-Bass.
- Schonlau, M., & Couper, M. P. (2016). Semi-Automated Categorization of Open-Ended Questions. *Survey Research Methods*, 10(2), 143-152.
- Schuman, H. (1966). The Random Probe: A Technique for Evaluating the Validity of Closed Questions. *American Sociological Review*, 31(2), 218-222.
- Schuman, H. (2008). *Method and Meaning in Polls and Surveys*. Cambridge, MA, Harvard University Press.
- Schuman, H., & Presser, S. (1977). Question Wording as an Independent Variable in Survey Analysis. *Sociological Methods and Research*, 6, 151-170.
- Schuman, H., & Presser, S. (1979). The Open and Closed Question. *American Sociological Review*, 44(5), 692-712.
- Singer, E. (2003). Exploring the Meaning of Consent: Participation in Research and Beliefs about Risks and Benefits. *Journal of Official Statistics*, 19, 273-286.
- Singer, E. (2011). Toward a Benefit-Cost Theory of Survey Participation: Evidence, Further Tests, and Implications. *Journal of Official Statistics*, 27(2), 379-392.
- Singer, E., & Couper, M. P. (2011). Ethical Considerations in Web Surveys. In M. Das, P. Ester, & L. Kaczmirek (eds.), *Social Research and the Internet* (pp. 133-162). New York: Taylor and Francis.
- Singer, E., & Kulka, R. A. (2002). Paying Respondents for Survey Participation. In *Studies of Welfare Populations: Data Collection and Research Issues*, edited by M. Ver Ploeg, R. A. Moffitt, & C. F. Citro. Washington, D.C.: National Academy Press.
- Smyth, J. D., Dillman, D. A., Christian, L. M., & McBride, M. (2009). Open-ended Questions in Web Surveys; Can Increasing the Size of Answer Boxes and Providing Extra Verbal Instructions Improve Response Quality? *Public Opinion Quarterly*, 73(2), 325-337.
- Sturgis, P., & Campanelli, P. (1998). The Scope for Reducing Refusals in Household Surveys: An Investigation Based on Transcripts of Tape-Recorded Doorstep Interactions. *Journal of the Market Research Society*, 40(2), 121-39.
- Tan, L. (2011). An Introduction to the Contact History Instrument (CHI) for the Consumer Expenditure Survey. Washington, DC: Bureau Labor Statistics, Consumer Expenditure Survey Anthology, 2011, pp. 8-16, available at <http://www.bls.gov/cex/anthology11/csxanth2.pdf>
- Thissen, M. R. (2014). Computer Audio-Recorded Interviewing as a Tool for Survey Research. *Social Science Computer Review*, 32(1), 90-104.

- Thissen, M. R., Fisher, C., Barber, L., & Sattaluri, S. (2008). Computer Audio-Recorded Interviewing (CARI), A Tool for Monitoring Field Interviewers and Improving Field Data Collection. Proceedings of the International Methodology Symposium 2008, Statistics Canada, Gatineau, Canada.
- Thissen, M. R., Park, H., & Nguyen, M. (2013). Computer Audio Recording: A Practical Technology for Improving Survey Quality. *Survey Practice*, 6(2).
- Thissen, M. R., & Rodriguez, G. (2004). Recording Interview Sound Bites Through Blaise Instruments. Paper presented at the International Blaise User's Conference, Gatineau, Québec, Canada, September.
- Tourangeau, R., Conrad, F. G., Couper, M. P., & Ye, C. (2014). The Effects of Providing Examples in Survey Questions. *Public Opinion Quarterly*, 78(1), 100-125.
- Tourangeau, R., Sun, H., Conrad, F. G., & Couper, M. P. (2016). Examples in Open-Ended Survey Questions. *International Journal Public Opinion Research*, advance online access, DOI: 10.1093/ijpor/edw015.
- Tourangeau, R., & Smith, T. W. (1996). Asking Sensitive Questions: The Impact of Data Collection Mode, Question Format, and Question Content. *Public Opinion Quarterly*, 60(2), 275-304.
- Wenemark, M. (2010). *The Respondent's Perspective in Health-Related Surveys*. Linköping University Medical Dissertations No. 1193.
- West, B. T. (2013). An Examination of the Quality and Utility of Interviewer Observations in the National Survey of Family Growth. *Journal of the Royal Statistical Society, Series A*, 176(1), 211-225.
- West, B. T., & Kreuter, F. (2013). Factors Affecting the Accuracy of Interviewer Observations: Evidence from the National Survey of Family Growth (NSFG). *Public Opinion Quarterly*, 77(2), 522-548.
- West, B. T., & Kreuter, F. (2015). A Practical Technique for Improving the Accuracy of Interviewer Observations of Respondent Characteristics. *Field Methods*, 27(2), 144-162.
- Yan, T., Curtin, R., & Jans, M. (2010). Trends in Income Nonresponse over Two Decades. *Journal of Official Statistics*, 26(1), 145-164.